系列决策任务中的策略转换：来自爱荷华赌博任务的证据*

胡馨允 沈悦 戴俊毅*
junyidai@zju.edu.cn
（浙江大学心理与行为科学系 杭州 310058）

摘要：

已有大量研究使用系列决策任务探讨了各类决策的决策策略。通过假定个体采用单一策略完成所有任务试次，并比较对应的计算认知模型拟合实证数据的能力，这些研究发现各种决策任务都涉及多种可能的决策策略。但是，此类研究的一个共同缺陷在于忽视了个体在任务过程中转换决策策略的可能性。通过开发允许在强化学习策略和启发式策略间转换的针对爱荷华赌博任务的计算认知模型，并将此类模型同单一策略模型进行对比，研究 1 提供了个体在该系列决策任务中会改变决策策略的明确证据。研究 2 则发现，随着试次数的增加，发生策略转换的可能性也会上升。这些结果表明，为了正确认识各种决策任务的决策策略，需要充分考虑在系列决策任务过程中发生策略转换的可能性，尤其是试次较多的系列任务。未来研究可以探讨策略转换的多种可能形式，以及导致策略转换的任务和个体因素，以便进一步深化对于系列决策任务的心理机制的认识。

关键词：系列决策任务，爱荷华赌博任务，策略转换，计算认知建模，强化学习和启发式策略

分类号：B841.2，B842.1

Strategy switching in a sequence of decisions: Evidence from the Iowa Gambling Task

Hu Xinyun, Shen Yue, Dai Junyi
(Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310058, China)

Abstract:

Much research has been devoted to studying decision strategies in various tasks. Such research usually involved a sequence of decision trials under the same task structure to provide sufficient information for inferring the underlying decision strategies. By assuming each individual adopted a single decision strategy across all decision trials and comparing corresponding computational cognitive models in terms of their performances in fitting empirical data, such studies have revealed multiple possible decision strategies for many major decision tasks. One common drawback of such research, however, was overlooking the possibility that individuals switched their strategies along the sequence of decisions. This might lead to inappropriate conclusions regarding the decision strategies underlying specific decision tasks or misleading inferences of potential cognitive and affective differences between normal and

* 本文系中央高校基本科研业务费专项资金资助项目（项目编号：2018QNA3014）的研究成果之一。

different clinical populations based on parameter estimates from models assuming single strategies.

To address this critical issue, two studies were conducted to examine the possibility of strategy switching in the Iowa Gambling Task (IGT), an experience-based decision task with a sequence of trials aimed at mimicking real-world decisions under uncertainty. By developing a computational cognitive model that allowed for switches between reinforcement learning strategies and heuristic strategies and comparing its performance with those of single-strategy models, Study 1 showed that data from about half of the 617 healthy participants in 10 previous studies were better fitted by the strategy-switching model than three single-strategy models that performed well in previous research, that is, the WSLS, PVL2, and VPP models as exemplar models assuming heuristic, reinforcement learning, and mixed strategies, respectively. This result provided clear support for the possibility of strategy switching in the IGT.

Since strategy switching might occur with accumulating experience or fatigue and an increasing number of trials is likely to facilitate such changes, 321 participants were recruited in Study 2 to further examine whether a larger number of trials would contribute to more strategy switching in the IGT. Specifically, 160 participants performed a 100-trial IGT, whereas the other 161 participants performed a 200-trial IGT under otherwise the same task structure. It was found that data from a larger proportion of individual participants were best fitted by the strategy-switching model when the IGT involved 200 trials rather than standard 100 trials. This result provided further evidence for strategy switching in the task.

Overall, the current results suggest that strategy switching is likely to occur in a sequence of decisions under the same task structure. Consequently, in order to obtain proper understanding of the decision strategies for various decision tasks, it is necessary to consider seriously the possibility of strategy switching, especially for a long sequence of decisions. For a more refined understanding of psychological mechanisms underlying sequences of decisions, future research might further investigate various forms of strategy switching such as gradual instead of abrupt switches and task and individual factors that trigger such switches.

Keywords: Decision task with a sequence of trials, The Iowa Gambling Task, Strategy switching, Computational cognitive modeling, Reinforcement learning and heuristic strategies

1 引言

古人云“明者因时而变，知者随事而制”，当重复面对任务结构相同的决策（即完成系列决策任务）时，人们所使用的决策策略不是一成不变的。[1]大量研究

[1] 本文探讨的系列决策任务有别于序列决策任务，后者一般是指后续决策的方案集合取决于之前的决策及其结果，即时间上相邻的决策存在明显的动态依存性的决策任务。

表明，各种决策任务都存在多种不同的决策策略。例如，针对多属性决策任务，存在一系列不同的补偿式（选项在不同属性上的优势和劣势可以相互抵消）和非补偿式策略（选项在不同属性上的优势和劣势不可相互抵消，例如， Payne et al.， 1988; Rieskamp & Otto, 2006; Walsh & Gluck, 2016)，而面对风险决策任务时，个体则可能采取基于期望效用或类似评估的策略 (例如， Kahneman & Tversky, 1979; Von Neumann & Morgenstern, 1944) 或者更为简单的启发式策略 (例如，Brandstätter et al.，2006)。此外，研究者还对信息环境、任务要求以及个体差异等因素如何影响个体的策略选择进行了探索 (例如， Bergert & Nosofsky, 2007; Pachur & Galesic, 2013)，并且发现，任务环境或者要求的变化可能会带来相应的决策策略的转换 (例如， Bröder & Schiffer, 2006; Lee et al.， 2014)。

除了由任务环境和要求的变化所导致的策略转换以外，人们是否还可能在相对稳定的任务环境和要求下，由于自我调整、适应或者内在的探索动机而发生策略转换？在绝大多数有关决策策略的实证研究中，被试都需要在相同的任务结构下完成一系列决策试次，以便研究者能够依托足够多的信息，来推断被试的决策策略。虽然过往研究已经探讨了面对特定决策任务时个体所使用的策略的多样性，以及影响策略选择的可能因素，却鲜有研究考察，在面对一个相对稳定的系列决策任务时，个体的决策策略发生转换的可能性。如果这种可能性的确存在，那么以往有关决策策略的研究，就会因为忽视这一可能性而导致错误的结论。为了更好地探明个体在面对各种决策任务时的决策策略，首先需要回答的问题是，在任务环境和要求相对稳定的系列决策中，是否的确会发生策略转换。本文将以爱荷华赌博任务这一典型的系列决策任务为例，探讨这一重要的理论和实践问题。

爱荷华赌博任务 (Iowa Gambling Task, IGT) 是一项基于经验的模拟决策任务，它最初是为了考察腹内侧前额叶损伤患者在应对不确定的现实情境时的决策缺陷而提出的 (Bechara et al.， 1994)。该任务包含四个牌堆（分别标记为A，B，C，D），被试需要多次在这些牌堆间做出选择。每次选择某一牌堆之后，都会抽取并翻转其最上方的一张牌，并根据牌面信息给予被试一定的奖励。但是，有时选择某一牌堆也会同时给被试带来损失。在任务开始之前，被试并不知道每个牌堆的盈亏规律以及总试次数，而他们的目标则是通过他们的选择获得尽可能高的总回报。因此，被试需要通过不断选择各个牌堆来学习每个牌堆的盈亏规律，并采取特定策略来完成这一任务。目前 IGT 已被广泛用于识别各种临床人群的决策缺陷，包括脑损伤人群 (Hochman et al.， 2010)、药物滥用人群 (Ahn et al.， 2014; Bechara & Damasio, 2002; Bechara et al.， 2001)、神经疾病人群 (Stout et al.， 2001) 以及精神障碍人群（李蕾等，2019；徐四华，2012）等。

除了被用于考察临床人群的决策缺陷，IGT 还被用来探究正常和临床人群在面对不确定情境时的决策策略。为此，研究者们提出了对应不同策略的一系列计算认知模型，这些模型大致可分为强化学习模型和启发式模型两类。强化学习模型假设 IGT 包含三个过程：涉及动机的对每次选择结果的评估过程，涉及认知的对牌堆期望效价的更新过程，以及涉及反应的概率化选择过程。Busemeyer 和 Stout (2002) 提出了第一个针对 IGT 的强化学习模型——期望效价学习 (Expectancy-Valence Learning, EVL) 模型。该模型假定个体使用期望效用 (Expectancy Utility, EU) 函数来评估每次选择结果的效用 (Ahn et al.， 2008)，使用差异学习 (Delta-Learning, DEL) 规则来更新每个牌堆的期望效

价（Rescorla & Wagner, 1972），并使用依赖于试次的选择（Trial-Dependent Choice, TDC）规则来指导下一试次的选择（Luce, 1959）。在 EVL 模型的基础上，Ahn 等人（2008）进一步探索了强化学习模型涉及的三个过程中每个过程的不同数学形式，并提出了预期效价学习（Prospect-Valence Learning, PVL）模型。该模型假定个体会使用预期效用（Prospect Utility, PU）函数（Kahneman & Tversky, 1979）对选择的净结果（即奖励以及可能同时出现的损失之和）进行评估，使用 Erev 和 Roth（1998）提出的衰减强化学习（Decay-Reinforcement Learning, DRL）规则更新预期效价，并且使用不随试次变化的选择（Trial-Independent Choice, TIC）规则（Yechiam & Ert, 2007）做出反应。更为近期的采用系统化模型比较方法的研究表明（Dai et al., 2015），个体在对结果进行评估时，更有可能会对同时出现的奖励和损失首先分别按照预期效用函数进行评估，然后再将评估结果加以整合。对应的模型被称为第 2 类预期效价学习（Prospect-Valence Learning 2, PVL2）模型。

在有关 IGT 的启发式模型中，最有代表性且拟合实证数据表现最好的是赢留输走（Win-Stay-Lose-Shift, WSLS）模型（Worthy et al., 2012）。该模型假设，人们的每次选择仅取决于上一次选择的牌堆以及所得的结果，而与更早之前的选择及其结果无关。因此，相比于考虑之前所有试次的选择及对应结果的强化学习模型，WSLS 模型假设的心理机制更为简单。具体而言，该模型假定个体继续选择相同牌堆的概率，受当前选择该牌堆的结果而定。如果当前选择的净结果非负（即赢），则有较大可能继续选择相同牌堆，反之（即输），则有较大可能下一试次转而选择不同的牌堆。

尽管关于 IGT 的决策策略已经有了丰富的研究成果，但很少有研究考虑个体在完成 IGT 过程中发生策略转换这一可能。Busemeyer 和 Stout（2002）曾提出过一个策略转换启发式选择（Strategy-Switching Heuristic Choice）模型。但是，该模型所谓的“策略转换”，并非是指决策策略的本质变化，而是指随着个体由于选择不利牌堆（即 A 或 B 牌堆）遭受越来越多的损失，其选择概率在不利牌堆和有利牌堆（即 C 或 D 牌堆）之间重新分配的过程。此外，也有研究者提出了将强化学习和启发式策略结合在一起的计算认知模型。例如，Worthy 等人（2013）提出了效价附加坚持（Valence-Plus-Perseverance, VPP）模型。该模型认为，在 IGT 的每一个试次中，人们都会综合考虑各个牌堆的期望效价以及前一试次的选择及其结果，再决定当前试次的选择。虽然该模型同时包含强化学习和启发式策略成分，且相比于 EVL、PVL 以及 WSLS 模型，该模型在拟合实证数据时有较好的表现，但它仍然假定个体会使用单一的，虽然更为复杂的混合策略来完成 IGT 中每个试次的选择。

综上所述，有关 IGT 的决策策略研究，尚未考察在任务过程中发生策略转换这一可能。如果个体的确会在任务过程中因为各种原因转变决策策略，那么以往仅仅比较单一策略模型的研究，就可能得出关于个体策略选择的错误认识。此外，那些根据单一策略模型的参数估计，来推断不同人群决策差异背后的心理机制的研究（例如，Ahn et al., 2014; Yechiam et al., 2005），也可能会产生有偏的估计，进而导致对人群差异的错误解读。本研究将通过开发允许策略转换的模型并将其与传统的单一策略模型进行比较，来回答在 IGT 中是否存在策略转换这一问题，以期为得出有关 IGT 中的决策策略以及不同人群差异的更为可信的结论提供依据，也为在更大范围内探讨决策策略转换这一重要的理论和实践问题提供借鉴。

2 研究一：IGT 策略转换模型的提出和检验

2.1 方法

2.1(1) IGT 简介

如上所述，IGT 包含四个牌堆（分别标记为 A、B、C、D），在每个试次中被试需要选择一个牌堆，并根据其最上方的牌呈现的信息获得一定的奖励，并有可能同时遭受一些损失。被试的目标是在总试次数未知的情况下，使总回报最大化。例如，在 Bechara 等人 (1994) 最早的 IGT 研究中包含了（被试未知的）100 个试次，并且采用了如表 1 所示的支付方案。具体而言，被试每次选择 A 或 B 牌堆，都会获得 100 美元的收益。但是，每选择 10 次 A 牌堆，被试都会遭受 5 次损失，金额从小到大分别为 150 美元、200 美元、250 美元、300 美元和 350 美元，且这 5 次损失在每 10 次选择中出现的具体位置都会有所变化。类似的，被试每选择 10 次 B 牌堆，都会遭受 1 次金额为 1250 美元的损失，且每 10 次选择中出现损失的位置也各不相同。对于 C 或者 D 牌堆，每次选择都会带来 50 美元的收益。然而，每选择 10 次 C 牌堆，都会遭受 5 次总额为 250 美元的损失，每选择 10 次 D 牌堆，则会遭受 1 次 250 美元的损失，且每 10 次选择 C 或 D 牌堆遭受损失试次的位置也会有所不同。后续研究使用了相同或者类似的任务设置，主要的调整发生在试次数，以及是否使用真实回报两方面。当使用真实回报（即按照被试最后的总回报支付酬金）时，出于控制实验经费的目的，一般会将 Bechara 等人最初的支付方案中的各种结果金额都缩减 100 倍 (例如, Dai et al., 2015)。无论采取何种支付方案，所有类型的 IGT 研究都满足以下三点：1）A 和 B 牌堆每次选择都有较高的收益，但总损失也较大，因此长期而言是不利的，即总回报为负；2）C 和 D 牌堆每次选择的收益较低，但总损失较小，因此长期而言是有利的，即总回报为正；3）A 和 C 牌堆相比于 B 和 D 牌堆会出现更多次的损失。

表 1 Bechara 等人 (1994) 使用的 IGT 支付方案

牌堆	A	B	C	D
每次选择的收益	100	100	50	50
每 10 次选择出现损失的次数	5	1	5	1
可能损失的金额	- 150 - 200 - 250 - 300 - 350	- 1250	- 25 - 50 - 75	- 250
10 次选择的总回报	- 250	- 250	250	250

2.1(2) 单一策略模型

为了给探究 IGT 中的策略转换提供合适的对照模型，本研究考虑了已有文献中的三大类单一策略模型，即强化学习模型，启发式模型以及混合模型，并以 PVL2 模型，WSLS 模型和 VPP 模型作为各类模型的代表。这些模型在以往的研究中都有较好的表现，因此如果新的允许策略转换的模型能够比它们有更好的表现，则能为 IGT 中存在策略转换提供支持。以下将介绍这三个计算认知模型的具体数学形式。

针对 IGT 的强化学习模型假定人们通过结果评估、期望（或预期）效价更新和概率化选择三个过程来完成该任务。根据 PVL2 模型 (Dai et al., 2015)，人们在选择某一牌堆之后，会针对当前选择获得的收益和可能的损失，使用预期理论的价值函数分别进行评估，然后再做汇总。其对应的效用函数被称为第 2 类预期效用 (Prospect Utility 2, PU2)函数，效用评估的具体形式如下：

$$u(t) = [win(t)]^{\alpha} - \gamma[|loss(t)|]^{\alpha} \quad (1)$$

其中，$win(t)$ 和 $loss(t)$ 分别代表在试次 t 获得的收益及可能同时出现的损失金额，$u(t)$ 代表试次 t 的汇总效用评估。α 是形状参数，用于衡量被试感受到的效用对于客观价值的敏感性，取值范围在 0 到 1 之间，γ 则代表预期理论中的损失厌恶参数，取值范围在 0 到 5 之间。

在完成了结果评估之后，根据 PVL2 模型，个体会使用衰减强化学习规则对各牌堆的预期效价进行更新，具体形式如下：

$$E_j(t) = A \cdot E_j(t-1) + \delta_j(t) \cdot u(t) \quad (2)$$

其中，$E_j(t)$ 代表牌堆 j 在第 t 个试次完成后的预期效价（j = 1, 2, 3, 4，分别对应于 A, B, C, D 四个牌堆），A 是记忆衰减参数，取值范围是 0 到 1，A 越大，表示记忆衰减对于预期效价的影响越小，$\delta_j(t)$ 是一个哑变量，当被试在试次 t 选择了牌堆 j 时为 1，否则为 0。换而言之，被选择的牌堆的预期效价更新既涉及记忆衰减，又涉及当前效用评估，而未选择牌堆的预期效价更新则只存在记忆衰减过程。

最后，PVL2 模型假定，个体会依据各牌堆的预期效价，使用以下函数确定下一次选择各牌堆的概率并相应地做出随机选择 (Sutton & Barto, 1998)：

$$\Pr[D(t+1) = j] = \frac{e^{\theta(t)\cdot E_j(t)}}{\sum_{k=1}^{4} e^{\theta(t)\cdot E_k(t)}} \quad (3)$$

其中，等式左侧的$\Pr[D(t+1) = j]$ 表示被试在试次 $t+1$ 选择牌堆 j 的概率，而等式右侧的分母是一个归一化因子，可以确保预测的各牌堆的选择概率之和为 1。$\theta(t)$ 是选择函数的灵敏度参数，$\theta(t)$ 越大，表明被试的选择越取决于牌堆的预期效价。根据 PVL2 模型， $\theta(t)$ 的取值不随试次的变化而变化，即 $\theta(t)$ 是 t 的常函数，其形式为：

$$\theta(t) = \theta = 3^{c} - 1 \quad (4)$$

其中，c 是自由参数，取值范围为 0 到 5。c 越大，θ 越大，代表被试更有可能选择预期效价较高的牌堆。总的来说，PVL2 模型包含了效用评估、预期效价更新和概率化选择三个过程，一共包含 α, γ, A 和 c 四个自由参数。

作为启发式模型的代表，WSLS 模型假定的决策策略比 PVL2 模型假定的策略明显更为简单。根据该模型，个体只会根据上一次选择的牌堆及其净结果（即收

益和损失的总和），来概率性地决定下一次的选择。该模型有两个参数，第一个参数代表上一次选择的牌堆得到的净结果大于等于 0 时，个体继续选择该牌堆的概率，即

$$\Pr(stay \mid win) = \Pr\left[D_j(t) \mid choice_{t-1} = D_j \,\&\, x(t-1) \geq 0\right] \quad (5)$$

其中 $choice_{t-1} = D_j$ 表示在 $t-1$ 试次选择了 j 牌堆，$x(t-1) \geq 0$ 表示该选择带来的净结果非负，而 $D_j(t)$ 则表示在 t 试次继续选择 j 牌堆。该模型的第二个参数代表上一次选择的牌堆净结果为负时，被试转而选择其他牌堆的概率，即

$$\Pr(shift \mid loss) = 1 - \Pr\left[D_j(t) \mid choice_{t-1} = D_j \,\&\, x(t-1) < 0\right] \quad (6)$$

其中符号的含义与公式 5 相同。该模型进一步假定，当在某一试次选择不同于上一试次的其他牌堆时，所有可能牌堆的选择概率相同。因此，为了保证选择所有牌堆的总概率为 1，当试次 $t-1$ 的净结果非负时，在试次 t 选择任意一个其他牌堆的概率为 $\frac{1-\Pr(stay \mid win)}{3}$；当试次 $t-1$ 的净结果为负时，在试次 t 继续选择相同牌堆的概率为 $1-\Pr(shift|loss)$，选择任意一个其他牌堆的概率则为 $\frac{\Pr(shift \mid loss)}{3}$。

除了强化学习模型和启发式模型，Worthy 等人 (2013) 提出的混合策略 VPP 模型也有很好的表现。Worthy 等人认为，使用衰减强化规则的强化学习模型混淆了坚持选择同一牌堆的倾向和选择预期效价最高的牌堆的倾向。因此，他们分离了这两种倾向，并提出了 VPP 模型。根据该模型，个体一方面会使用 PU 函数来对某次选择结果进行效用评估，并使用差异学习规则更新牌堆的预期效价，其具体形式如下：

$$u(t) = \begin{cases} x(t)^{\alpha} & \text{当 } x(t) \geq 0 \\ -\lambda|x(t)|^{\alpha} & \text{当 } x(t) < 0 \end{cases} \quad (7)$$

$$E_j(t) = E_j(t-1) + A \cdot \delta_j(t) \cdot [u(t) - E_j(t-1)] \quad (8)$$

其中，$x(t)$ 表示当前试次选择结果的净收益，其他符号的含义同上文。

另一方面，个体还会根据之前试次是否选择了牌堆 j 以及选择牌堆 j 所得净收益是否非负来确定当前试次坚持选择牌堆 j 的倾向，具体形式如下：

$$P_i(t) = \begin{cases} k * P_i(t-1) + \epsilon_{pos} & \text{当 } x(t) \geq 0 \\ k * P_i(t-1) + \epsilon_{neg} & \text{当 } x(t) < 0 \end{cases} \quad (9)$$

其中 $P_j(t)$ 代表在试次 t 坚持选择牌堆 j 的倾向，k 是一个取值范围在 0 到 1 之间的衰减参数，其含义类似于公式 2 中的参数 A，ϵ_{pos} 和 ϵ_{neg} 是两个自由参数，代表一个牌堆被选择后，由其净结果决定的坚持倾向的改变量，范围

都在 - 1 到 1 之间。最终，每个牌堆的价值（$V_j(t)$，即综合评价）是其预期效价和坚持倾向的加权平均，具体形式如下：

$$V_j(t) = w_{E_j} \cdot E_j(t) + \left(1 - w_{E_j}\right) \cdot P_j(t) \tag{10}$$

其中，w_{E_j} 是权重参数，取值范围在 0 到 1 之间。

最后，和 PVL2 模型类似，VPP 模型假设被试会根据牌堆的价值确定下一次选择各牌堆的概率并相应地做出随机选择，具体规则如下：

$$\Pr[D(t+1) = j] = \frac{e^{\theta(t) \cdot V_j(t)}}{\sum_{k=1}^{4} e^{\theta(t) \cdot V_k(t)}} \tag{11}$$

其中，灵敏度参数 $\theta(t)$ 的计算使用的是不随试次变化的规则，即公式 4。VPP 模型共包含八个自由参数，即涉及效用评估和更新的 α, λ, A，涉及坚持程度的 $k, \epsilon_{pos}, \epsilon_{neg}$，以及涉及综合评价的 w_{E_j} 和涉及选择反应的 c。

2.1(3) 策略转换模型

由于 IGT 一般包含多达 100 个甚至更多的试次，在整个任务过程中，个体可能由于各种原因发生策略转换。在本研究中，我们假定可能存在两种转换，一种是在任务开始阶段由于缺乏信息而使用对信息依赖度较低的启发式策略，并在对各牌堆有了更多了解之后，转而使用更为复杂更为精细的强化学习策略。另一种则是在初始阶段就使用强化学习策略，并随着任务的进行，因为疲劳、倦怠或者降低认知负荷的需求，转而采用启发式策略。从建模角度，鉴于 PVL2 模型在强化学习模型，以及 WSLS 模型在启发式模型中的优势地位，本研究将分别以这两个模型来表达可能的强化学习策略和启发式策略，并由此探讨个体在 IGT 中发生策略转换的可能性。

具体而言，我们开发了一个允许发生一次策略转换 (Switching-Strategy-Once, SSO) 的模型。该模型假设个体在完成 IGT 的过程中，会在启发式策略和强化学习策略之间进行一次转换，且个体在使用启发式或者强化学习策略完成 IGT 时所使用的具体计算认知机制，和对应的 WSLS 或者 PVL2 模型所假定的机制相同。除了 WSLS 模型和 PVL2 模型涉及的参数以外，该模型还包含两个新的参数，分别代表发生策略转换的节点试次，记作 sp（即 Switching Point），以及策略转换的类型，记作 st（即 Switching Type）。$st = 1$ 代表个体在完成 IGT 的过程中先使用了强化学习策略，之后转而使用启发式策略，而 $st = 2$ 则代表相反的策略转换过程。因此，该模型共有 8 个参数，即涉及强化学习策略的 α, γ, A 和 c，涉及启发式策略的 $\Pr(stay|win)$ 和 $\Pr(shift|loss)$，转换节点参数 sp，以及转换类型参数 st。由于当策略转换节点位于整个任务的开始或结尾阶段时，相应的策略转换模型和对应的单一策略模型可能过于类似，难以分辨。因此，在本研究中，我们将 sp 的范围限定在第 21 个试次到倒数第 21 个试次之间。

2.1(4) 数据

为了系统比较策略转换模型和单一策略模型拟合实证数据的能力，我们选取了以往采用 IGT 的研究中具有代表性的一系列数据集作为模型拟合对象 (Steingroever et al., 2015)。具体而言，这些数据出自 10 项研究，涵盖了不同年龄范围的共 617 名健康被试，且 IGT 的试次数包含 95，100 和 150 三种情况。所有研究中的 IGT 都在计算机上完成，且支付方案与表 1 所示的 Bechara 等人 (1994) 所用的方案相同或类似。所涉及的各项研究的基本信息参见 Steingroever 等人的表 1。

2.1(5) 模型拟合和比较方法

本研究所考察的每个计算认知模型（即 WSLS, PVL2, VPP 和 SSO），都可以根据被试之前的选择以及所得结果，预测下一试次每个牌堆被选择的概率（即一步向前预测，Ahn et al., 2008）。因此，我们首先使用极大似然估计法 (Maximum-Likelihood Estimation, MLE)，用每个模型去拟合个体被试的选择数据，即找到每个模型下，可以使得实际选择数据出现可能性最大化的参数取值组合，并以相应的观测数据的预测出现概率，作为模型拟合表现的初步指标。具体而言，在特定模型参数取值下的似然值被定义为该取值下，模型预测的个体被试的选择序列的发生概率，而对数似然值 (Log-Likelihood, LL) 则被定义为

$$LL = \sum_{t=1}^{n-1}\sum_{j=1}^{4} \ln\left(\Pr\left(D_j(t+1)\right)\right) \times \delta_j(t+1) \qquad (12)$$

其中，n 表示总试次数，$\Pr\left(D_j(t+1)\right)$ 表示模型基于被试前 t 次选择及其结果，所预测的第 $t+1$ 试次选择牌堆 j 的概率。$\delta_j(t+1)$ 是一个哑变量，如果被试在 $t+1$ 试次选择了牌堆 j，则 $\delta_j(t+1)=1$，否则 $\delta_j(t+1)=0$。这意味着，每个试次只有实际被选择的牌堆的预测概率会被纳入对数似然值的计算之中。然后，我们使用 MATLAB 中的 PSO 算法 (Particle Swarm Optimization, Clerc, 2010) 来寻找每个模型对数似然值的最大值，并求得对应的参数估计值。

一般而言，更为复杂的模型会有更好的拟合表现。由于上述模型的参数个数不尽相同，它们的复杂程度也不尽相同。因此，我们使用包含二阶偏差修正的赤池信息准则 (Akaike Information Criterion with second-order bias correction, AIC_C; Akaike, 1974; Sugiura, 1978) 和贝叶斯信息准则 (Bayesian Information Criterion, BIC; Schwarz, 1978) 这两种常用的适用于极大似然估计的指标，来综合考量模型的拟合情况和复杂程度，并以相应的准则分数来评价每个模型的表现并进行模型选择，具体计算方式如下：

$$AIC_C = -2LL_M + 2k + \frac{2k(k+1)}{n-k-1} \qquad (13)$$

$$BIC = -2LL_M + k \times \ln(n) \qquad (14)$$

其中，k 代表模型的自由参数个数，n为需要拟合的数据点个数（即总试次数-1），而 LL_M 则是指模型的极大对数似然值。AIC$_C$（或 BIC）的值越小，表示模型表现越好 (Broomell et al., 2011)。[2]

2.1(6) 模型复原测试

由于 AIC$_C$ 和 BIC 对于拟合表现所做的调整存在程度上的差异，且一般而言 BIC 的惩罚程度（即 $k*\ln(n)$）要高于 AIC$_C$的惩罚程度（即 $2k+\frac{2k(k+1)}{n-k-1}$），所以使用这两种指标可能会导致不同的模型选择结果。因此，我们还进行了模型复原测试，以便确定哪一指标更适用于针对观测数据进行模型选择(Wagenmakers et al., 2004; Worthy et al., 2012)。具体而言，该测试有以下两个主要步骤：第一，针对每个模型，使用拟合观测数据得到的最优参数取值产生模拟的被试数据。在模拟过程中，不会使用到被试实际的选择及其结果，而是根据模型预测的选择概率以及 IGT 本身的设定，来随机地产生模拟数据；第二，用不同模型拟合每种模型产生的模拟被试数据，并采用特定模型选择指标来比较模型表现。如果使用某种模型选择指标时，每个模型都只针对自身产生的数据有相对较好的表现，那么说明该选择指标下，模型的区分度较大。相反，如果使用某种模型选择指标时，某些模型针对别的模型产生的数据也会有相对较好的表现，则说明该选择指标下，模型的区分度不大。换而言之，这样的选择指标不能较为准确地确认出产生数据的真实模型，因此也就不适用于根据观测数据进行模型选择。

在本研究中，我们对数据集中的 617 名被试的观测数据进行了模型拟合，从而得到了每个被试在每个模型下的最优拟合参数取值。然后，对于每个模型，我们用对应于每名被试的最优拟合参数取值产生 3 组模拟数据，共产生 1821（=617×3）组模拟的被试数据。之后，我们分别使用 WSLS 模型、PVL2 模型、VPP 模型和 SSO 模型，用拟合观测数据一样的方法拟合这些模拟数据。最后，通过分析使用不同指标（即 AIC$_C$ 和 BIC）时模型的区分度，我们可以选取出更为合理的针对观测数据的模型选择指标。

2.2 结果

2.2(1) 模型拟合和比较

表 2 展示了各个模型拟合全部 617 名被试的观测数据的结果。当以 AIC$_C$ 为模型选择指标时，无论是就群体均值还是个体结果而言，SSO 模型都表现最佳，而 VPP、PVL2 和 WSLS 模型的表现则依次变差。当以 BIC 为模型选择指标时，就群体均值而言，PVL2 模型的表现最佳，SSO 模型次之。从个体结果上看，WSLS 模型和 PVL2 模型表现较好，分别在 30.79%和 33.87%的被试数据上有最好的表现，而 VPP 和 SSO 模型的表现则基本相当。无论采用 AIC$_C$ 还是 BIC 作为指标，SSO 模型都在一部分被试的数据（AIC$_C$：43.27%，BIC：18.96%）上有最好的表现。

表 2 研究一模型比较结果

模型	以 AIC$_C$ 为指标	以 BIC 为指标

[2]当样本量与模型参数个数的比值较小（即样本量/参数个数<40）时，使用包含二阶偏差修正的赤池信息准则 (AIC$_C$) 能够弥补使用 AIC 可能导致的过拟合缺陷 (Burnham & Anderson, 2004)。因此，在本文中我们使用 AIC$_C$ 而非 AIC 作为模型评估的一个指标。

	指标均值（标准差）	该模型表现最好的被试人数及比例	指标均值（标准差）	该模型表现最好的被试人数及比例
WSLS	236.27 (72.65)	42（6.81%）	241.46 (72.76)	190（30.79%）
PVL2	225.25 (72.28)	114（18.48%）	235.48 (72.49)	209（33.87%）
VPP	220.37 (70.25)	194（31.44%）	240.13 (70.71)	101（16.37%）
SSO	219.60 (71.06)	267（43.27%）	239.36 (71.48)	117（18.96%）

2.2(2) 模型复原测试

由于 AIC_C 和 BIC 对于模型复杂度的惩罚程度存在差异，相比于 BIC，AIC_C 倾向于选择参数更多的模型。因此，出现使用 AIC_C 指标时，较为复杂的 VPP 和 SSO 模型有更好的表现并不奇怪。为了选择更合适的模型选择指标，我们进行了模型复原测试。表 3 和表 4 展示了模型复原测试的结果。当以 AIC_C 为模型选择指标时，各模型有较好的区分度。对于每个模型产生的模拟被试数据，该模型本身都能在最大比例的个体模拟数据上有最好的表现。而当以 BIC 为模型选择指标时，对于每个模型产生的模拟数据，最为简单的 WSLS 模型都能在最大比例的个体模拟数据上有最好的表现，即 BIC 不能很好地对 WSLS 和其他模型进行区分。因此，在本研究中，相比于 BIC，将 AIC_C 作为模型选择指标更为合适。

表 3 研究一基于 AIC_C 的模型复原测试结果

数据产生模型	数据拟合模型			
	WSLS	PVL2	VPP	SSO
WSLS	88.60%	3.67%	0.92%	6.81%
PVL2	33.55%	46.14%	10.97%	9.35%
VPP	14.37%	16.69%	59.97%	8.97%
SSO	13.99%	7.73%	2.76%	75.53%

注：表中的每一行代表不同模型在某个模型产生的模拟被试数据上的表现情况。例如，第一行代表各个模型拟合 WSLS 模型产生的模拟被试数据时的表现。在由 WSLS 模型产生的模拟被试数据中，WSLS 模型在 88.60%的个体数据上表现最佳，而 PVL2 模型、VPP 模型和 SSO 模型则分别在 3.67%、0.92%和 6.81%的个体数据上表现最佳。

表 4 研究一基于 BIC 的模型复原测试结果

数据产生模型	数据拟合模型			
	WSLS	PVL2	VPP	SSO
WSLS	99.57%	0.43%	0.00%	0.00%
PVL2	51.43%	44.79%	3.62%	0.16%
VPP	37.39%	32.63%	29.07%	0.92%
SSO	42.63%	20.31%	0.05%	37.01%

注：表中内容的含义同表 3。

2.3 讨论

本研究提出了有关 IGT 的一次策略转换模型，并针对以往 617 名健康被试的数据，比较了此模型和假定单一策略的具有代表性的 PVL2 模型（强化学习策略），WSLS 模型（启发式策略）以及 VPP 模型（混合策略）的数据拟合表现。当分别以 AIC_c 和 BIC 作为模型选择指标时，模型表现的相对优劣有所差异，但策略转换模型都能在一定比例的个体数据上有最好的表现。模型复原测试的结果表明，AIC_c 比 BIC 更适合在当前研究中被用于进行模型选择，因为相比于使用 BIC，在使用 AIC_c 时更可能还原出正确的数据产生模型。当以 AIC_c 作为模型选择指标时，SSO 模型无论从群体还是个体水平都要优于另外三个模型，而且策略转换模型在近一半（43.27%）的被试观测数据上表现最佳。这些结果表明，个体在完成 IGT 的过程中，的确有较大可能会发生决策策略的转换。

如前所述，经验累积或者疲倦等因素可能是造成在像 IGT 这样的系列决策任务中发生策略转换的原因。当任务的试次数变得越来越多时，我们可以合理地认为，经验累积或者疲倦这样的因素更有可能发生作用，因而个体也就更有可能在任务过程中，变换决策策略。因此，作为本研究主体部分的补充，我们还比较了包含不同试次数的 IGT 研究中的模型表现，以便进一步考察策略转换的可能性。在本研究考察的 617 名被试中，有 15 人完成的是 95 试次的 IGT，504 人完成的是 100 试次的 IGT，还有 98 人完成的是 150 试次的 IGT。表 5 展示了包含不同试次数的 IGT 数据以 AIC_c 为模型选择指标的相应结果。可以看出，随着试次数的上升，无论是从 AIC_c 均值，还是从模型表现最好的被试比例来看，策略转换模型相比于其他模型的优势都在增强，这一点在模型表现最好的个体被试比例上表现得尤为明显，即从 13.33%上升到了 53.06%。

表 5 研究一中根据试次数分组的模型拟合和比较结果

模型	AIC_c 均值（标准差）			该模型表现最好的被试人数及比例		
	95 试次	100 试次	150 试次	95 试次	100 试次	150 试次
WSLS	238.66 (35.45)	224.42 (56.54)	296.81 (111.01)	1（6.67%）	36（7.14%）	5（5.10%）
PVL2	223.20 (39.06)	215.21 (58.10)	277.19 (110.48)	5（33.33%）	94（18.65%）	15 (15.31%)
VPP	222.84 (40.43)	210.29 (55.93)	271.84 (108.07)	7（46.67%）	161（31.94%）	26 (26.53%)
SSO	227.14 (37.92)	210.14 (57.76)	267.10 (108.68)	2（13.33%）	213（42.26%）	52 (53.06%)

为了更加深入地了解策略转换的具体情况，我们进一步分析了不同试次数下，策略转换节点（即 sp）的分布状况。具体而言，针对各种试次数的 IGT 任务，我们计算了 SSO 模型拟合得最好的个体数据对应的 sp 参数的分布信息。当总试次数为 95 时，sp 估计值的均值为 48.5，标准差为 37.48；当总试次数为 100 时，sp 估计值的均值为 48.92，标准差为 19.47，而当总试次数为 150 时，sp 估计值的均值为 81.42，标准差为 36.03。不难发现，随着 IGT 试次数的增多，发生策略转换的平均位置也在后移。针对每种转换类型（即从强化学习策略转化为启发式策略，$st = 1$，或者从启发式策略转化为强化学习策略，$st = 2$），我们进一步使用单侧 Mann-Whitney 检验分析了 100 试次和 150 试次下的平均转换节点的差异（在完成 95 试次 IGT 的被试中，仅有 2 人的数据 SSO 模型拟合得最好，故此处不做分析）。结果发现，无论是先使用强化学习策略，还是先使用启发式

策略的被试，当需要完成 150 试次 IGT 时，相应的平均转换节点，均显著晚于只需完成 100 试次 IGT 时的平均转换节点（p 值均小于 0.001）。[3]之所以会出现这一状况，可能是由于随着 IGT 试次数的增多，有更高比例的被试在整个任务过程中发生了策略转换，且新增的发生策略转换的被试的转换节点较晚。这为个体在完成 IGT 过程中有可能发生策略转换，且随着试次数的增多，被试会有更大的可能性发生策略转换提供了进一步的证据。

需要指出的是，虽然上述分析支持 IGT 中可能存在策略转换，但这些分析所考察的数据出自不同的研究，在任务设置的细节上不尽相同，而且试次数的范围和间距不尽合理，完成不同试次数 IGT 的人数也很不均衡。因此，以上分析结果只能被认为是为支持 IGT 中的策略转换提供了有限的证据。在以下报告的研究二中，我们在对试次数进行更为合理的操纵的前提下，采用相同的任务设置在每种试次数下收集了人数几乎相同的被试数据，以便更好地检验试次数增加会提升策略转换的可能性这一关键假设。

3 研究二：试次数对 IGT 中策略转换可能性的影响

3.1 方法

3.1(1) 被试

本研究采用实验范式操纵 IGT 的试次数，并设置了 100 试次和 200 试次两个实验条件。共招募 321 名成年大学生被试（男性 134 人，女性 187 人），平均年龄 20.54 岁（SD=2.41）。其中 160 人完成了 100 试次的 IGT，另 161 人则完成了 200 试次的 IGT。招募被试时要求非心理学专业且未参加过 IGT 研究。所有被试均在实验前填写知情同意书，并自愿参与实验。实验结束后，被试会得到基础报酬和额外奖励，额外奖励的数量和 IGT 的绩效有关，绩效越高，额外奖励越多。

3.1(2) 实验设计与流程

本实验采用单因素被试间设计，考察并比较不同试次数下个体在 IGT 中发生策略转换的可能性。本实验共设置 100 试次和 200 试次两种实验条件，前者是大多数 IGT 研究的标准设置，而后者则可以在控制实验总时长的前提下，有效地拉开与前者的距离，以实现一定程度的效应量。

任务开始前，被试会阅读有关 IGT 的标准化介绍，并被告知拥有 2000 元研究货币（即初始总财富）。任务开始后，被试会看到分别位于屏幕上、下、左、右侧的四个牌堆，并可以通过键盘的“上”、“下”、“左”、“右”键，选择对应的牌堆。被试在完成任务之前，并不知晓所需完成的试次数。每次选择完成后，屏幕中央将呈现当前试次的奖励和损失，以及更新之后的总财富额（如图 1）。设置以上下左右方式呈现牌堆，是为了减少传统的从左到右的排布方式对牌堆选择产生的非随机的影响，例如在开始阶段依次选择 A、B、C、D 四个牌堆，以及在后续试次中，相继选择空间上明显相邻的牌堆。此外，本研究采用和表 1 所示相同的支付方案，且每 10 次选择某一牌堆时损失出现的试次位置也是随机的。

[3] 我们也使用单侧 Mann-Whitney 检验考察了不同转换类型下的转换节点差异，发现仅在 100 试次条件下两种转换类型间存在显著差异。在后续的研究 2 中，无论是 100 试次 IGT 还是 200 试次 IGT，平均转换节点在不同转换类型间都不存在显著的差异。

实验程序使用 Python3 及 PsychoPy 软件编写，被试需要在电脑的 PsychoPy 软件上完成实验。

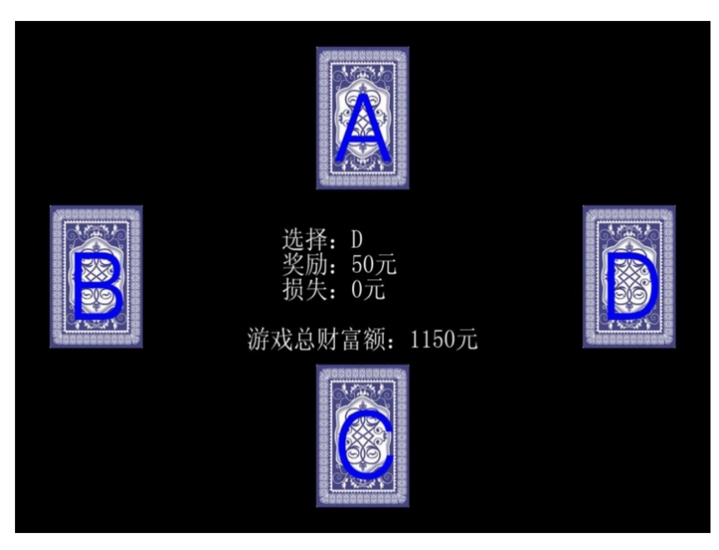


图 1 研究二实验界面截图

3.1(3) 数据分析

本研究采用和研究一相同的模型拟合和比较技术，分析和比较了 3 个单一策略模型和一次策略转换模型在拟合个体 IGT 数据时的表现，并且进行了模型复原测试。此外，使用独立样本比例差异 Z 检验，分析试次数对于 IGT 中发生策略转换的可能性的影响。

3.2 结果

3.2(1) 模型分析和比较

因模型复原测试表明，在本研究中使用 AIC_C 仍然比使用 BIC 更有可能做出正确的模型选择（见下文），此处仅报告基于 AIC_C 的结果。表 6 呈现了以 AIC_C 为标准，100 和 200 试次组各自的模型比较结果。无论是从群体均值，还是从个体结果来看，SSO 模型在两种试次数条件下都表现最佳。而且，无论是针对 100 试次 IGT 还是 200 试次 IGT，SSO 模型都在至少一半被试的个体数据上有最好的表现。此外，和研究一一样，VPP、PVL2 和 WSLS 模型的表现依次变差。独立样本比例差异 Z 检验的结果表明，200 试次下发生策略转换的可能性（即 SSO 模型在拟合个体观测数据时表现最佳的比例，65.22%），高于 100 试次下发生策略转换的可能性(50.00%，z=2.76，单侧 p=0.003，比例差异的 95% CI = [0.045, 0.259]，Cohen’s h = 0.31，对应较小的效应量）。

表 6 研究二模型比较结果

模型	AIC_C 均值（标准差）		各模型表现最好的被试人数及比例	
	100 试次	200 试次	100 试次	200 试次
WSLS	221.29 (54.27)	413.04 (125.04)	17 (10.63%)	4 (2.48 %)
PVL2	214.36 (56.19)	392.14 (123.53)	27 (16.88%)	15 (9.32%)
VPP	212.66 (52.58)	383.47 (120.92)	36 (22.50%)	37 (22.98%)
SSO	207.95 (54.46)	377.02 (120.76)	80 (50.00%)	105 (65.22%)

和在研究一中一样，我们还分析了两种试次数条件下，SSO 模型拟合最优的那些被试的 sp 参数的估计结果。当 IGT 包含 100 试次时，sp 估计值的均值为 47.03，标准差为 20.39；当 IGT 包含 200 试次时，sp 估计值的均值为 95.38，标准差为 54.21。[4]单侧 Mann-Whitney 检验结果表明，无论在哪种转换类型下，200 试次下的平均转换节点均显著晚于 100 试次下的平均转换节点（p 值均小于 0.001）。

3.2(2) 模型复原测试

本研究使用每个模型模拟了 3×321=963 组个体被试数据，并使用 4 个模型对每组模拟数据进行了拟合。表 7 展示了 100 试次组和 200 试次组基于 AIC_C 的模型复原测试结果。不论是在 100 试次还是 200 试次下，所考察的每个模型都能在最大比例的各自模型产生的模拟数据上有最好的表现。总体而言，试次数为 200 时数据生成模型被正确复原的比例（71.74%），要高于试次数为 100 时的比例（64.69%，z = 4.70, 单侧 p < 0.001, 比例差异的 95% CI = [0.041, 0.100], Cohen’s h = 0.15, 对应小的效应量）。

表 7 研究二基于 AIC_C 的模型复原测试结果

	WSLS	PVL2	VPP	SSO
WSLS	90.83%/86.13%	2.29%/3.11%	1.04%/0.83%	5.83%/9.94%
PVL2	38.54%/27.33%	44.17%/53.21%	3.13%/5.18%	14.17%/14.29%
VPP	23.75%/13.66%	17.29%/12.63%	46.46%/63.77%	12.50%/9.94%
SSO	15.21%/10.14%	5.21%/4.35%	2.29%/1.66%	77.29%/83.85%

注：每个单元格中的前一个数值代表 100 试次组的结果，后一个数值代表 200 试次组的结果。

表 8 展示了基于 BIC 的模型复原测试结果。可以看出，和研究一一样，当使用 BIC 进行模型选择时，几乎在所有情况下，无论针对哪个模型产生的个体模拟数据，WSLS 模型都能有最好的表现，即 BIC 不能很好地对 WSLS 和其他模型进行区分。只有当试次数为 200 时，PVL2 模型和 SSO 模型才能在各自产生的模拟数据上有最好的表现。总体而言，试次数为 200 时数据生成模型被正确复原的比例（59.06%），要高于试次数为 100 时的比例（49.17%, z = 6.16, 单侧 p < 0.001, 比例差异的 95% CI = [0.068, 0.130], Cohen’s h = 0.20, 对应小的效应量）。

表 8 研究二基于 BIC 的模型复原测试结果

	WSLS	PVL2	VPP	SSO
WSLS	100.00%/100.00%	0.00%/0.00%	0.00%/0.00%	0.00%/0.00%
PVL2	59.58%/46.38%	39.79%/53.42%	0.00%/0.00%	0.63%/0.21%

[4] 在本研究以及研究 1 中，SSO 模型拟合最优的被试的 sp 平均估计值都接近于允许范围的中间值。造成这一结果的可能原因是，发生策略转换的个体的策略转换节点位于模型允许范围内的各个位置的可能性大致相当，且整体分布呈单峰形态。

VPP	48.54%/35.20%	27.50%/34.16%	23.33%/29.81%	0.63%/0.83%
SSO	47.71%/32.30%	18.13%/14.70%	0.63%/0.00%	33.54%/53.00%

注：表中内容的含义同表 7。

3.3 讨论

本研究的目的在于考察试次数的增加是否会导致被试在 IGT 中更有可能发生策略转换。结果表明，无论 IGT 包含标准的 100 个试次还是更多的 200 个试次，和研究一一样，策略转换模型都在至少一半被试的个体数据上有最好的表现。更为重要的是，同包含 100 个试次的 IGT 相比，当 IGT 包含 200 个试次时，策略转换模型在更高比例的个体数据上表现最佳。这意味着，当试次数为 200 时，人们更有可能在 IGT 中发生策略转换。这一结果排除了策略转换模型能够在部分被试的数据上有最好的表现，仅仅是由模型比较结果的随机性所致这一解释，从而为个体在像 IGT 这样的系列决策任务中可能发生策略转换提供了进一步的支持。此外，模型复原测试的结果表明，与 BIC 相比，AIC_C仍然是更有可能做出正确的模型选择的指标。因此，本研究继续使用 AIC_C作为模型选择和策略推断的依据。最后，无论是采用 AIC_C还是 BIC 作为模型选择指标，200 试次下的模型复原表现，都要优于 100 试次下的表现。这与更大的数据量将有助于更好地区分不同模型的传统看法是一致的。

4 总讨论

系列决策任务既广泛存在于我们的日常生活中，也大量出现在有关决策策略和影响因素的实证研究之中。例如，为了招聘各种岗位的职员，人力资源部门的员工需要频繁地在求职者间做出选择，而像 IGT 这样的需要被试在相同的任务结构下重复完成多次决策的实验室任务也比比皆是。以往有关系列决策任务下的决策策略的研究，一般假设个体在所有试次中都使用相同的策略。之所以要求进行多次重复决策，仅仅是为了给推断决策策略提供更多的信息。但是，在这样的决策任务中，人们不仅会了解和学习任务刺激的具体特征，而且可能在更高的水平上，学习和相应地调整他们的决策策略。对于后一种学习的充分了解，将有助于我们得出有关策略选择的更为准确的推断，并且考察影响策略选择及其转换的因素，从而更好地为改善决策服务。

本研究以 IGT 为对象，较为系统地探讨了人们在系列决策任务中发生策略转换的可能性。结果表明，人们不仅会在 IGT 中发生策略转换，而且这一转换的可能性，还会随着任务试次数的上升而有所提升。这表明，在通过各种系列决策任务探讨个体的决策策略时，需要充分考虑策略转换的可能性，尤其是在任务试次数较多的情况下。具体而言，可以参照本文所报告的方式，开发允许策略转换的计算认知模型，并将它们和假定单一策略的模型进行比较，从而推断个体是否发生了策略转换，以及在何时发生了策略转换。由此，研究者有望对个体在任务不同阶段的策略使用情况有更加准确的认识，后续基于不同阶段的模型参数估计的分析，也更有可能产生相对准确的推断。

尽管本文报告的研究提供了有关个体在 IGT 中可能发生策略转换的明确证据，但这些研究所考虑的策略转换，仅是可能的多种策略转换类型中的一部分。具体而言，我们假定在整个任务过程中，个体只可能发生一次在强化学习策略和

启发式策略之间的转换，而且这种转换是以突变方式进行的。同样有可能出现的情况是，个体在任务过程中发生了多次策略转换，或者策略转换是以渐进的方式发生的，即在相继的试次中，从主要采取强化学习策略向主要使用启发式策略过渡，或者反向而行。从建模的角度，前一种可能性需要引入多个转换节点，而后一种则需要借助像 VPP 模型这样的混合模型，并假设其中有关不同策略的加权系数（即 w_{E_j}）是试次的渐变函数。就分析技术而言，实现这样的模型都更有挑战性，但是并非完全没有可能。例如，针对多属性系列决策任务，Lee 等人 (2019, 2021) 采用贝叶斯方法探索了允许发生多次策略转换的模型，发现有部分被试的数据能够被策略转换模型更好地解释，且有少量被试的数据支持存在多次策略转换。未来的研究可以参考 Lee 等人的方式，探讨 IGT 或者其他重要的决策任务（例如风险和跨期选择任务）下发生多次策略转换的可能性，还可以开发策略渐变模型，并将此类模型和（单次或多次）突变模型进行比较，从而加深对于策略转换的多种可能性的认识。除了最终的选择，与任务有关的其他数据也可能反映了人们的决策过程，例如决策反应时和眼动数据等。Fang 等人 (2023) 基于鼠标追踪技术获取的数据，提出了用于多属性决策任务的机器学习策略识别 (Machine Learning Strategy Identification, MLSI) 方法。这种通过使用机器学习算法提取决策特征并进而甄别决策策略的研究方法十分新颖，未来可以在有关策略转换的研究中进一步推广使用。需要指出的是，就核心决策策略而言，针对多属性决策的模型（例如，Take-the-Best 模型）一般都是确定性模型，而针对 IGT 的模型则一般是概率性模型，因此后者在数据分析层面更为复杂。此外，IGT 和多属性决策任务存在一些重要的区别。例如，前者的每次决策都有明确的反馈，且所做的选择和相应结果会对后续决策产生影响，因此更多涉及记忆和经验累积因素，而后者的每次决策则是相对独立的，且一般不包含反馈信息，因而无需记忆和经验的参与。此外，多属性决策任务一般是在确定信息条件下完成的，而 IGT 则涉及更为复杂的不确定信息。因此，来自多属性决策任务和 IGT 的策略转换证据属于不同类型任务下的汇聚证据。这些证据表明，策略转换可能存在于性质不同的各种系列决策任务之中。

在确认了系列决策任务存在策略转换的可能性后，一个需要进一步探讨的关键问题是，产生策略转换的条件是什么，或者说怎样的任务因素、个体因素或者两者的交互可能引发策略转换。例如，当任务难度或者自身的抱负水平较高时，个体可能因为现有策略无法实现目标，而选择尝试不同的策略。由此可以推断，通过增大任务难度（比如要求在 IGT 中必须使得财富水平有所增长）或者提升个体的抱负水平的方式，也许能够引发更多的策略转换。此外，是否存在优势策略也是影响策略转换的一个可能因素。当个体在尝试了不同策略并且发现了优势策略之后，其策略转换的倾向可能会有所减弱。反之，如果多种策略下的任务表现大致相当，那么发生策略转换的可能性则将取决于个体希望尽可能有更好的表现的意愿，以及探索不同策略的动机程度。对于策略转换诱发因素的考察，将进一步提升我们对于决策策略及其转换的认识。

参考文献

[1] Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., & Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive Science,*

32(8), 1376 - 1402. https://doi.org/10.1080/03640210802352992

[2] Ahn, W. Y., Vasilev, G., Lee, S. H., Busemeyer, J. R., Kruschke, J. K., Bechara, A., & Vassileva, J. (2014). Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. Frontiers in Psychology, 5, 849. https://doi.org/10.3389/fpsyg.2014.00849

[3] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716 - 723.

[4] Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1 - 3), 7 - 15. https://doi.org/10.1016/0010-0277(94)90018-3

[5] Bechara, A., & Damasio, H. (2002). Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. Neuropsychologia, 40(10), 1675 - 1689. https://doi.org/10.1016/s0028-3932(02)00015-5

[6] Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. Neuropsychologia, 39(4), 376 - 389. https://doi.org/10.1016/s0028-3932(00)00136-6

[7] Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. Journal of Experimental Psychology: Learning, Memory, & Cognition, 33, 107 - 129.

[8] Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: making choices without trade-offs. Psychological Review, 113, 409 - 432.

[9] Bröder, A., & Schiffer, S. (2006). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. Journal of Experimental Psychology: Learning, Memory, & Cognition, 32, 904 - 918. https://doi.org/10.1037/0278-7393.32.4.904

[10] Broomell, S. B., Budescu, D. V., & Por, H. H. (2011). Pair-wise comparisons of multiple models. Judgment & Decision Making, 6(8), 821 - 831.

[11] Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods & Research, 33(2), 261 - 304.

https://doi.org/10.1177/0049124104268644

[12] Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. Psychological assessment, 14(3), 253. https://doi.org/10.1037/1040-3590.14.3.253

[13] Clerc, M. (2010). Particle swarm optimization (Vol. 93). John Wiley & Sons.

[14] Dai, J., Kerestes, R., Upton, D. J., Busemeyer, J. R., & Stout, J. C. (2015). An improved cognitive model of the Iowa and Soochow Gambling Tasks with regard to model fitting performance and tests of parameter consistency. Frontiers in Psychology, 6, 299. https://doi.org/10.3389/fpsyg.2015.00229

[15] Erev, I., & Roth, A. E. (1998). Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. American Economic Review, 88(4), 848–881. https://jstor.org/stable/117009

[16] Fang, J., Schooler, L., & Shenghua, L. (2023). Machine learning strategy identification: A paradigm to uncover decision strategies with high fidelity. Behavior Research Methods, 55(1), 263-284.

[17] Hochman, G., Yechiam, E., & Bechara, A. (2010). Recency gets larger as lesions move from anterior to posterior locations within the ventromedial prefrontal cortex. Behavioural Brain Research, 213(1), 27–34. https://doi.org/10.1016/j.bbr.2010.04.023

[18] Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. Econometrica, 47(2), 263–292. https://jstor.org/stable/1914185

[19] Lee, M. D., & Gluck, K. A. (2021). Modeling strategy switches in multi-attribute decision making. Computational Brain & Behavior, 4, 148–163. https://doi.org/10.1007/s42113-020-00092-w

[20] Lee, M. D., Gluck, K. A., & Walsh, M. M. (2019). Understanding the complexity of simple decisions: Modeling multiple behaviors and switching strategies. Decision, 6(4), 335–368. https://doi.org/10.1037/dec0000105

[21] Lee, M. D., Newell, B. R., & Vandekerckhove, J. (2014). Modeling the adaptation of search termination in human decision making. Decision, 1(4), 223–251. https://doi.org/10.1037/dec0000019

[22] Li, L., Zhang, J. Q., Hou, J. W., Li, Y. L., Lu, Y. J., & Guo, Z. J. (2019). Decision-

making characteristics assessed by the IOWA Gambling Task in schizophrenia: A meta-analysis. Journal of Qingdao University (Medical Sciences), 55(6), 688 - 691, 695.
[李蕾, 张俊青, 侯继文, 李亚铃, 鲁玉洁, 郭宗君. (2019). 爱荷华赌博任务评估精神分裂症决策特点Meta分析. 青岛大学学报(医学版), 55(6), 688 - 691, 695.]

[23] Luce, R. D. (1959). Individual choice behavior: A theoretical analysis. New York: Wiley.

[24] Pachur, T., & Galesic, M. (2013). Strategy selection in risky choice: the impact of numeracy, affect, and cross-cultural differences. Journal of Behavioral Decision Making, 26, 260 - 271.

[25] Payne, J. W., Bettman, J. R., & Johnson. E. J. (1988). Adaptive strategy selection in decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 534 - 552.

[26] Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical conditioning II: current research and theory, 2, 64 - 99.

[27] Rieskamp, J., & Otto, P. E. (2006). SSL: a theory of how people learn to select strategies. Journal of Experimental Psychology: General, 135(2), 207 - 236. https://doi.org/10.1037/0096-3445.135.2.207

[28] Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 461-464.

[29] Steingroever, H., Fridberg, D. J., Horstmann, A., Kjome, K. L., Kumari, V., Lane, S. D., … Wagenmakers, E. J. (2015). Data from 617 healthy participants performing the Iowa Gambling Task: A “Many Labs” Collaboration. Journal of Open Psychology Data, 3(1), e5. http://doi.org/10.5334/jopd.ak

[30] Stout, J. C., Rodawalt, W. C., & Siemers, E. R. (2001). Risky decision making in Huntington's disease. Journal of the International Neuropsychological Society, 7(1), 92 - 101. https://doi.org/10.1017/s1355617701711095

[31] Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. Communications in Statistics, Theory and Methods A, 7, 13 - 26. http://doi.org/10.1080/03610927808827599

[32] Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: an introduction. IEEE

Transactions on Neural Networks, 9(5), 1054–1054. https://doi.org/10.1109/tnn.1998.712192

[33] Von Neumann, J., & Morgenstern, O. (1944). Theory of games and economic behavior. Princeton University Press.

[34] Wagenmakers, E. J., Ratcliff, R., Gomez, P., & Iverson, G. J. (2004). Assessing model mimicry using the parametric bootstrap. Journal of Mathematical Psychology, 48, 28–50. https://doi.org/10.1016/j.jmp.2003.11.004

[35] Walsh, M. M., & Gluck, K. A. (2016). Verbalization of decision strategies in multiple-cue probabilistic inference. Journal of Behavioral Decision Making, 29(1), 78–91. https://doi.org/10.1002/bdm.1878

[36] Worthy, D. A., Hawthorne, M. J., & Otto, A. R. (2012). Heterogeneity of strategy use in the Iowa gambling task: a comparison of win-stay/lose-shift and reinforcement learning models. Psychonomic Bulletin & Review, 20(2), 364–371. https://doi.org/10.3758/s13423-012-0324-9

[37] Worthy, D. A., Pang, B., & Byrne, K. A. (2013). Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task. Frontiers in Psychology, 4, 640. https://doi.org/10.3389/fpsyg.2013.00640

[38] Xu, S. H. (2012). Internet addicts' behavior impulsivity: evidence from the Iowa Gambling Task. Acta Psychologica Sinica, 44(11), 1523–1534.
[徐四华. (2012). 网络成瘾者的行为冲动性——来自爱荷华赌博任务的证据. 心理学报, 44(11), 1523–1534.]

[39] Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. Psychological Science, 16, 973–978.

[40] Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models. Journal of Mathematical Psychology, 51(2), 75–84. https://doi.org/10.1016/j.jmp.2006.11.002

(通讯作者：戴俊毅 E-mail: junyidai@zju.edu.cn)

作者贡献声明：

胡馨允：清洗和分析数据，论文起草；

沈悦：设计研究方案，采集、清洗和分析数据；

戴俊毅：提出研究思路，设计研究方案，论文最终版本修订。